

Abstract

- Accurate hydrologic modelling is critical to effective water resource management.
- Catchment attributes influence the hydrologic behaviors in an area and inform hydrologic models to better predict the discharge in a basin.
- The difficulty in predicting discharge may also be related to the complexity of the discharge signal.

Objectives

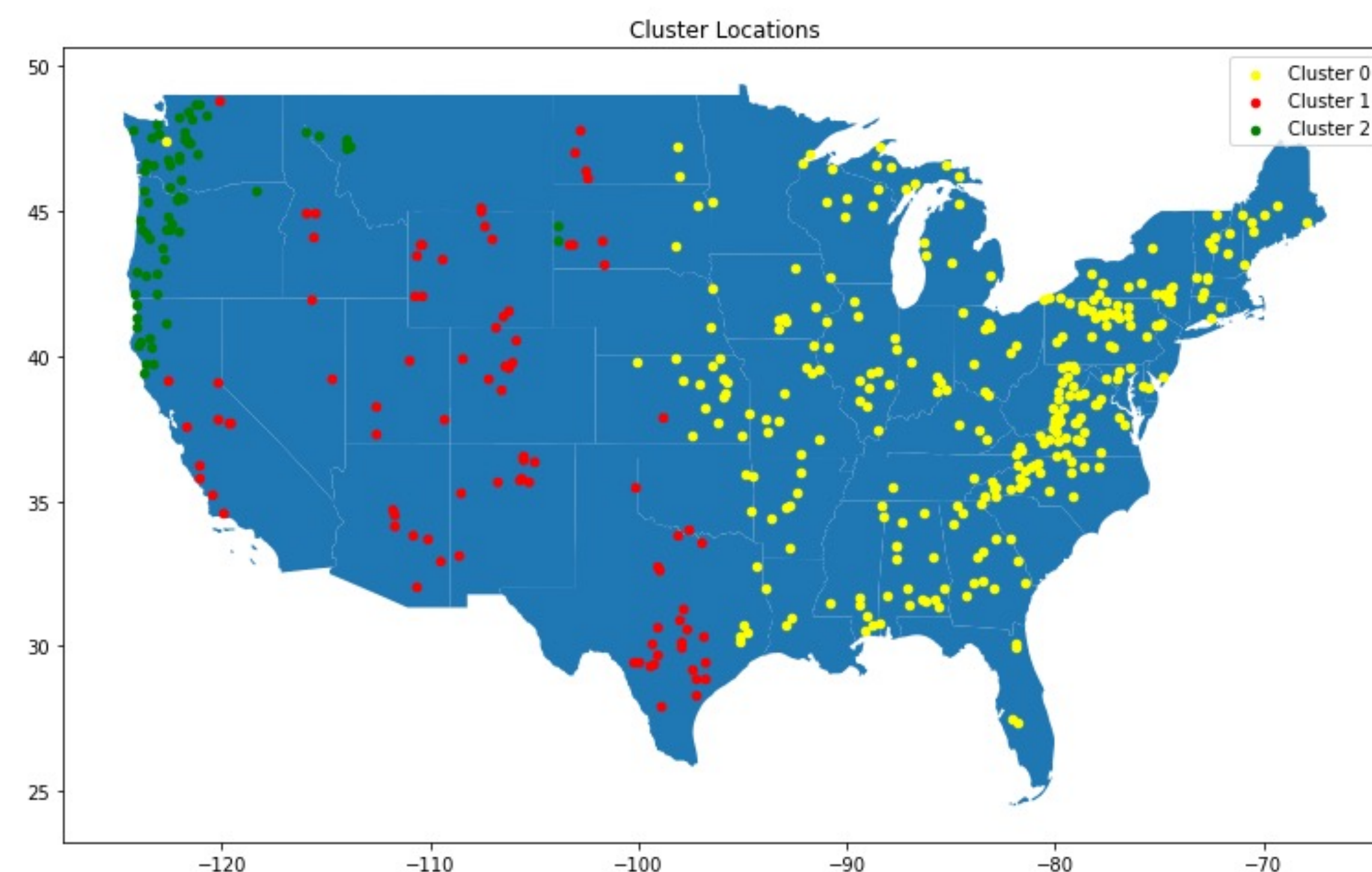
- Are there inherent properties to a catchment that make it easier to model?
- The study establishes the relationship between a catchment's static attributes and hydrologic model performance, and investigates the link to complexity, which we quantify with compressibility based in information theory.
- Hypothesis:** High model performance is mostly driven by low complexity hydrographs, which are expected to be connected to catchment attributes.

Approach

- A large national dataset and the performance metrics for corresponding hydrologic models are analyzed (obtained from [1, 2]).
- Group basins by similar catchment attributes using Hierarchical Clustering after filtering data with Principal Component Analysis, then look for performance and attribute trends within the clusters.
- Compressibility of basin hydrographs measured and trends compared [3].

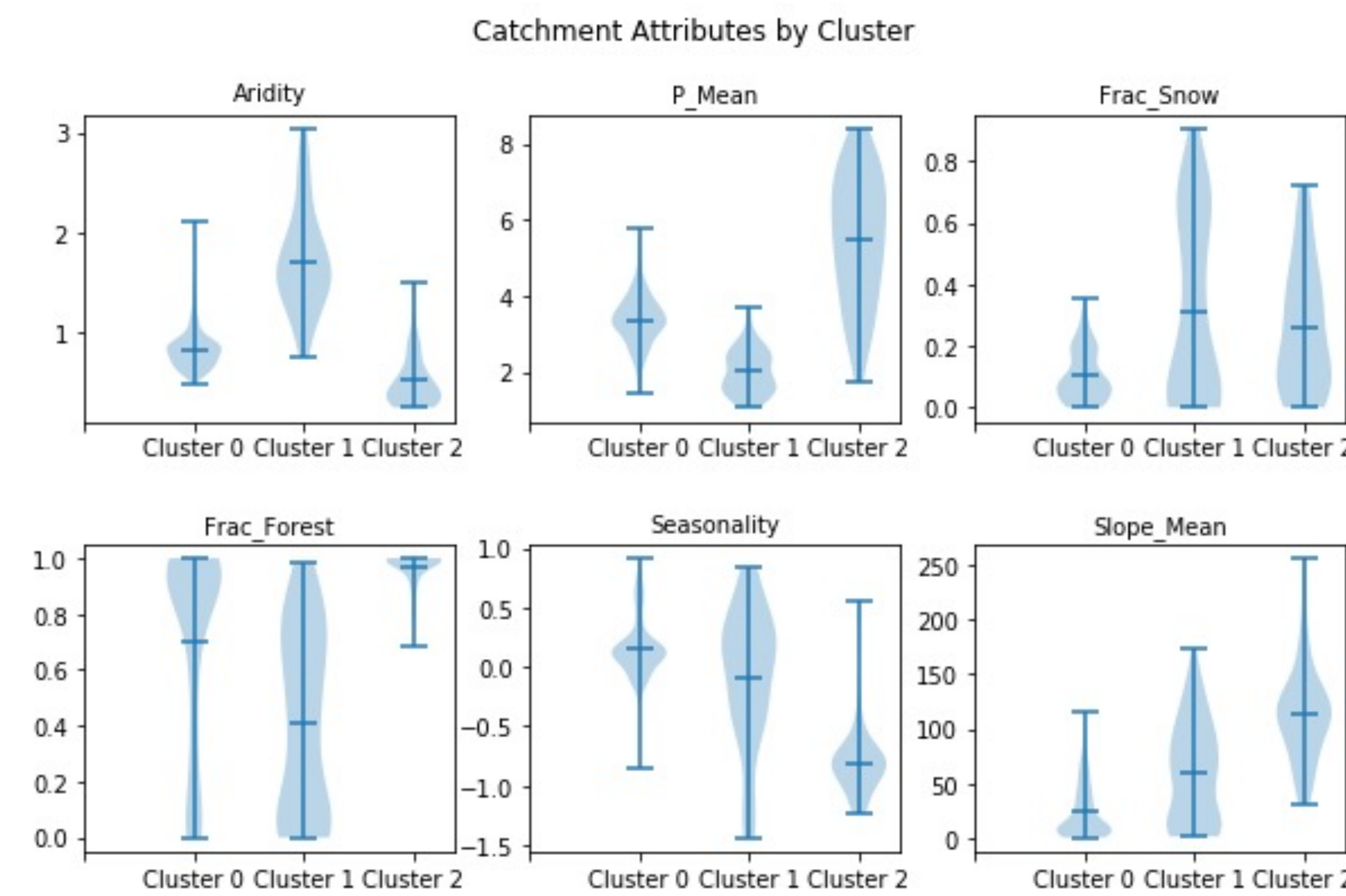
Cluster US Map

- Basins in the CONUS were clustered according to their similarities in catchment attributes.
- Three separate clusters were identified.



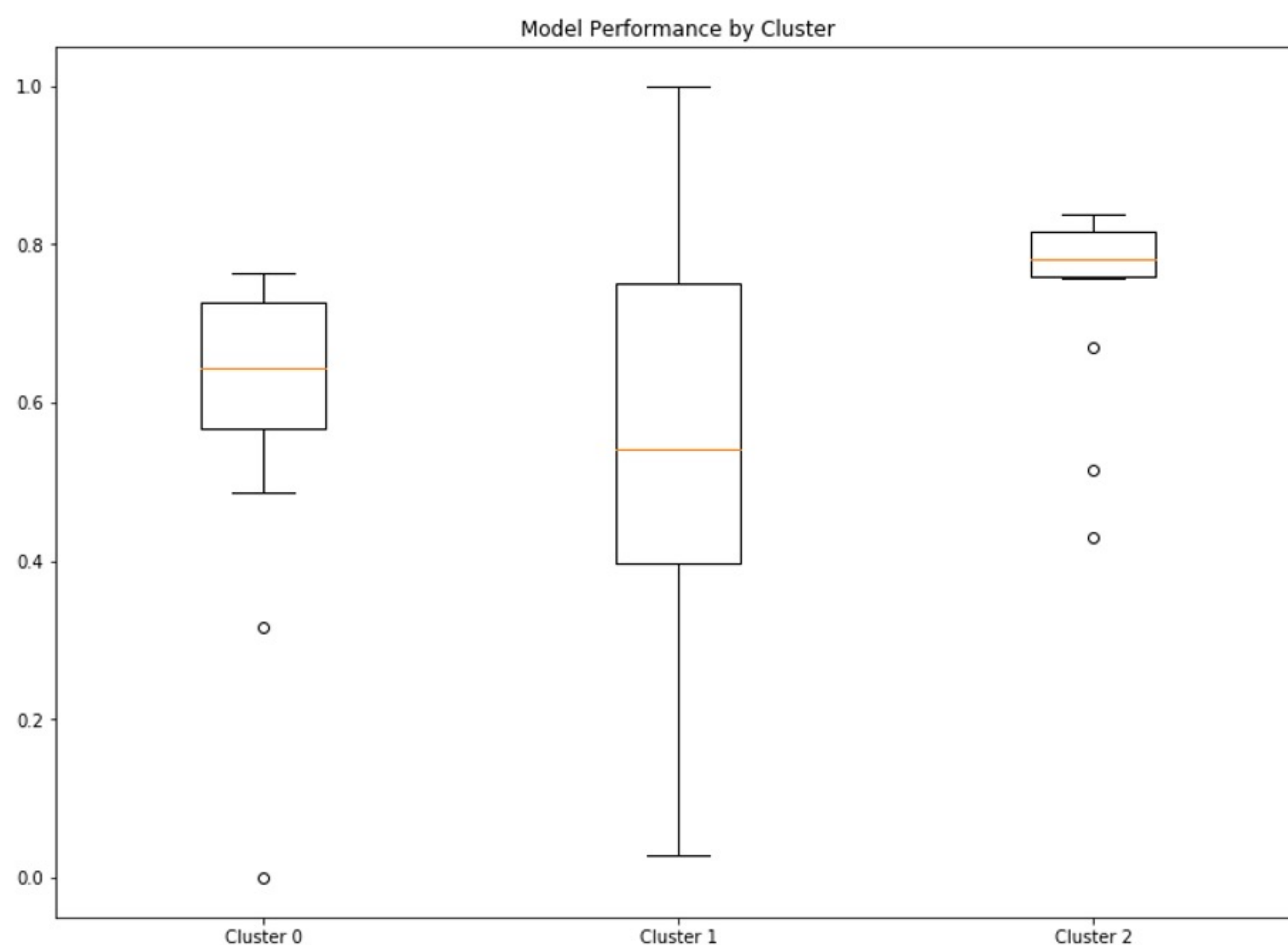
Cluster Characteristics

- Cluster 0 (289 basins): flat, summer dominated precipitation
- Cluster 1 (91 basins): arid, uniform annual precipitation
- Cluster 2 (67 basins): wet, forested, winter dominated precipitation



Cluster Performance

- Cluster 2, the forested cluster emerged with the strongest model performance, and least variance within the cluster.
- Cluster 1 performed the poorest, which aligns with the basin's high aridity.



Complexity, Predictability, and Catchment Characteristics

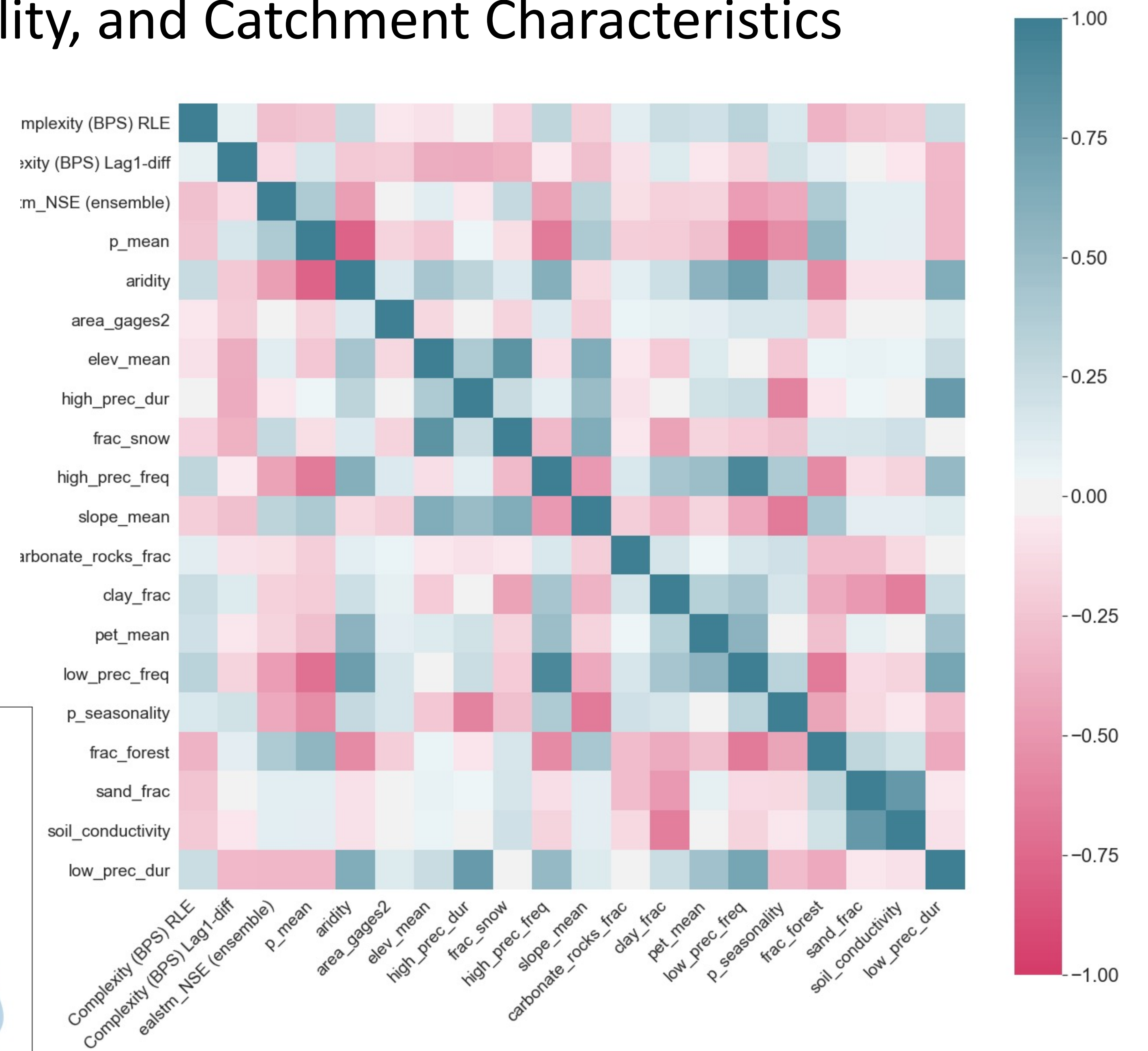
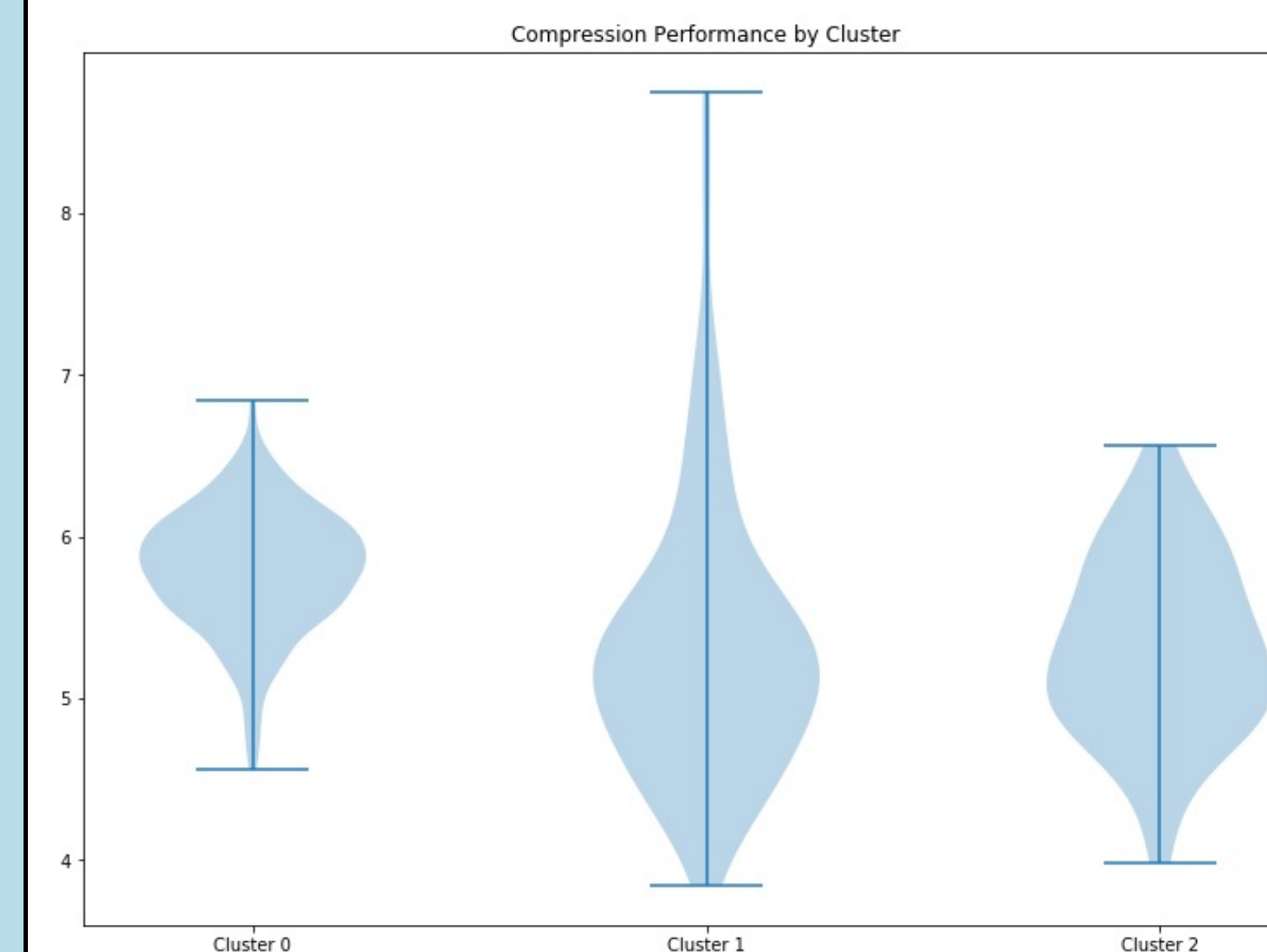
Method

The complexity analysis, using Hydrozip [3], only focuses on the temporal dependency aspect (temporal structure gives a more compressible signal).

Low bits per sample (BPS) means a compressible, low complexity signal.

Result

Correlations between predictability and complexity are relatively weak. Basins with strong influence of snow melt seem to show better compressibility.



Discussion

The weak relation between predictability and compressibility may indicate streamflow complexity stems mostly from precipitation complexity. Hydrological models can exploit the fact that streamflow has relatively low conditional complexity, given rainfall. Recent machine learning approaches are able to extract that structure. The choice of discretization (binning) of the timeseries in the compression experiment needs further investigation.

Conclusions

- Significant differences in model performance emerged between the clusters of basins.
- The compression analysis did not yield any substantial differences between the basin groups.
- Our initial hypothesis is most likely false, but needs more investigation into methodological aspects.

Future Work

- Investigate impact of different binning strategies on compression analysis results.
- Use a basin's catchment attributes in a predictive manner to hypothesize the model performance of new basins.

References

- [1] Addor, N., et al. (2017). Catchment Attributes for Large-Sample Studies. Hydrol. Earth Syst. Sci., 21, 5293–5313. Retrieved from: <https://doi.org/10.5065/D6G73C3Q>
- [2] Kratzert, F., et al (2019),. Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine Learning Applied to Large-Sample Datasets. Hydrology and Earth System Sciences, 23(12), 5089-5110.
- [3] Weijs, S.V. et al (2013). HydroZIP: How Hydrological Knowledge can Be Used to Improve Compression of Hydrological Data. Entropy 2013, 15, 1289-1310. <https://doi.org/10.3390/e15041289>

Acknowledgements

The authors are grateful to the Environment and Climate Change Canada, the British Columbia Graduate Scholarship, Sam Anderson, and all the people from the UBC HydroInfo Theory laboratory involved in the project, especially Dan Kovacek.